

# Universal interpretation of efficacy parameter in perturbed nonequilibrium systems

Sourabh Lahiri\* and A. M. Jayannavar†

*Institute of Physics, Sachivalaya Marg, Bhubaneswar - 751005, India*

The fluctuation theorems have remained one of the cornerstones in the study of systems that are driven far out of equilibrium, and they provide strong constraints on the fraction of trajectories that behave atypically in light of the second law. They have mainly been derived for a predetermined external drive applied to the system. However, to improve the efficiency of a process, one needs to incorporate protocols that are modified by receiving feedbacks about the recent state of the system, during its evolution. In such a case, the forms of the conventional fluctuation theorems get modified, the correction term involving terms that depend on the way the reverse/conjugate process is defined, namely, the rules of using feedback in order to generate the exact time-reversed/conjugate protocols. We show in this paper that this can be done in a large number of ways, and in each case we would get a different expression for the correction terms. This would in turn lead to several lower bounds on the mean work that must be performed on the system, or on the entropy changes. Here we analyze a form of the extended fluctuation theorems that involves the efficacy parameter, and find that this form gives rise to a lower bound for the mean work that retains a consistent physical meaning regardless of the design of feedback along the conjugate process, as opposed to the case of the previously mentioned form of the modified fluctuation theorems.

PACS numbers: 05.40.-a, 05.70.Ln, 05.20.-y

## I. INTRODUCTION

The theory of classical thermodynamics has been well supported by the methods of equilibrium statistical physics, which in turn is based on the Boltzmann probability distribution for the states of a system in phase space. For systems that are slightly away from equilibrium (where the perturbation to the system can be assumed to be linearly coupled to the applied weak force field), the linear response theory characterizes the response of the system in terms of equilibrium correlation functions. However, there are only very few exact results that remain valid even when the system is driven arbitrarily out of equilibrium. The fluctuation theorems have been one of them. First introduced for thermostatted systems following dissipative but deterministic dynamics by Evans and Searles [1, 2], it has produced several closely related equalities [3–8], some theorems holding in general for all systems while the others being satisfied for special cases (for example, in the limit of a large observation time). In a nutshell, these theorems have the following general form:

$$\frac{P(X_t)}{P^T(X_t^T)} = e^{X_t}, \quad (1)$$

where  $X_t$  is some observable that is to be measured and is in general a path function,  $P(X_t)$  is the probability density of this observable along a process (usually parametrized by an externally controlled time-dependent protocol  $\lambda(t)$ ) and  $P^T$  is the probability density whose

functional form is related to  $P$  through a conjugate transformation which is not necessarily time-reversal. Its argument  $X_t^T$  is the value assumed by the observable along the transformed trajectory in phase space. Some observables that have been shown to follow the fluctuation theorems are entropy, work and heat [3–10]. It is easy to see that the theorems are closely related to the second law, in the sense that the averaged observable  $\langle X_t \rangle$  always abides by the inequalities dictated by the second law. For instance, the total entropy of the system plus environment can be shown to give rise to the inequality [7, 8]

$$\langle \Delta s_{tot} \rangle \geq 0. \quad (2)$$

What this inequality says is that, even though some individual realizations of the experiment may observe a decrease in the total entropy, its *average* value must never decrease with time.

An issue that has recently attracted much interest is the use of feedback-controlled protocol to enhance the efficiency of such a process [11–20]. Instead of using a predetermined protocol, we can perform intermediate measurements on some observable and modify our protocol accordingly. We are in general not interested in the measuring device involved in the experiment (which definitely belongs to the “universe” in which second law must be obeyed). It is obvious that the form of the second law, now dealing with only a part of the universe under consideration, needs to be modified. As expected, the correction term involves the properties of the ignored subpart, in the form of the mutual information gained through measurements and gives rise to the following inequality:

$$\langle X_t \rangle \geq -\langle I \rangle. \quad (3)$$

\*Electronic address: lahiri@iopb.res.in

†Electronic address: jayan@iopb.res.in

Here, the mutual information is defined as [11–16]

$$I \equiv \ln \frac{p(m_1|x_1)p(m_2|x_2)\cdots P(m_N|x_N)}{P(m_1, m_2, \dots, m_N)}, \quad (4)$$

where the presence of measurement errors is assumed. In the arguments of the conditional probability  $p(m_i|x_i)$ ,  $m_i$  is the outcome when a measurement is performed, while  $x_i$  is the actual value of the observable. The averaging has been done over all possible phase space trajectories *as well as* over all possible protocol functions, which results due to the different feedback rules corresponding to different outcomes. Since  $\langle I \rangle$  is a relative entropy and consequently always non-negative [21], the above inequality provides us with the privilege of extracting work from or obtaining a decrease in entropy of the system, provided we can formulate a nice feedback algorithm. Note that eq. (3) only says that the lower bound for the dissipated work (or entropy) *can* be decreased by an amount proportional to the average mutual information. However, it does not provide any clue to what a good feedback algorithm should be. In this sense, it is reasonable to look for a quantity that actually provides a quantitative measure of how efficient an employed feedback procedure is. This is the so-called *efficacy parameter* [13] defined below. In this paper, we clarify the physical meaning and some of the properties of this parameter. The Jarzynski equality [3, 4] has two different generalized forms in the presence of feedback. The more commonly used form of the extended Jarzynski Equality (EJE) is [11–16]

$$\langle e^{-\beta W_d[X] - I[X, M]} \rangle = 1. \quad (5)$$

$W_d$  is the dissipated work defined through  $W_d = W - \Delta F$ , where  $\Delta F$  is the difference between free energies at the end and at the beginning of a particular protocol. Here, the reverse trajectories are generated by simply reversing the sequence of one of the forward protocols and we have used upper case letters to denote the full path of successive phase points or measured outcomes:  $X \equiv (x_0, x_1, \dots, x_\tau)$ ,  $M \equiv (m_1, m_2, \dots, m_N)$ . The time has been discretized as  $(t_0, t_1, t_2, \dots, t_N, t_{N+1})$ . We will set  $t_0 = 0$  and  $t_{N+1} = \tau$ . As we will show in this paper,  $I[X, M]$  will in general be replaced by a different physical quantity  $\phi[X, M]$ , if we choose to use feedback along the reverse process as well, as will be detailed below.

Other than eq. (5), there is yet another form of the extended Jarzynski Equality that has been introduced in the literature [12, 13]:

$$\langle e^{-\beta W_d[X]} \rangle = \gamma. \quad (6)$$

The efficacy parameter  $\gamma$  is a functional of the feedback control used along the forward process, and that determines the extent to which the feedback is efficient (i.e., more work can be extracted from the system). If  $\gamma = 1$ , then we would have the Jarzynski equality in absence of feedback. The efficacy parameter has been measured experimentally [22] and equations (5) and (6) have been

verified. Using the Jensen's inequality, we have the dissipated work bounded from below through the relation

$$\langle W_d[X] \rangle \geq -\ln \gamma \quad (7)$$

Although this inequality looks similar to the one stated before, it may seem to be a trivial statement because it is a simple consequence of the definition of  $\gamma$ . However, the fact that this definition lends a very clear physical meaning to the efficacy parameter and that this meaning can be exploited to experimentally measure  $\gamma$  (without using the definition eq. (6)) underlines the importance of the above bound. We explicitly show that  $\gamma$  retains the simple meaning when we extend eq. (6) to driven systems that make excursions among nonequilibrium steady states under arbitrary feedback-controlled protocols, whereas the physical meaning of  $\phi[X, M]$  depends on the design of feedback along the reverse process.

## II. EXTENDED JARZYNSKI EQUALITY

### A. Blind time-reversal of protocol

As shown in [11–16], in presence of information gain and feedback applied *only* along the forward trajectory, the Jarzynski Equality gets modified to eq. (5). This relation is easily derived from the detailed fluctuation theorem, in the case when the reverse protocol is the blind time-reversal of the corresponding protocol along the forward process. Then the ratio of the forward and reverse trajectories become [14–16]

$$\begin{aligned} \frac{P[X, M]}{\tilde{P}[\tilde{X}; M]} &= \frac{P[X]}{\tilde{P}[\tilde{X}]} \times \frac{p(m_1|x_1)p(m_2|x_2)\cdots P(m_N|x_N)}{P(m_1, m_2, \dots, m_N)}, \\ &= e^{\beta W_d[X] + I[X, M]}. \end{aligned} \quad (8)$$

In the above equation,  $\tilde{X}$  is the time-reversed trajectory  $(\tilde{x}_\tau, \tilde{x}_N, \dots, \tilde{x}_0)$ , and  $\tilde{P}[\tilde{X}]$  is the probability density for this time-reversed trajectory (for simplicity of notation, we would always use the forward time  $t$  even during the reverse process, in place of using the actual time  $\tilde{t} = \tau - t$  elapsed during the reverse process). If the state is denoted only by the position variable, then we have  $\tilde{x}_i = x_i$ .  $\tilde{P}[\tilde{X}; M]$  is the probability path functional for the *actual reverse process* being carried out. This is different from  $\tilde{P}[\tilde{X}, \tilde{M}]$  which gives the probability of obtaining the reverse phase space trajectory along with the reversed measured outcomes, provided measurements are performed at the same time instants as in the forward process. The expression for the latter quantity has no explicit reference to the actual feedback procedure followed to generate the reverse process (i.e., its physical meaning does not change due to change in feedback procedure for reverse process). A simple cross-multiplication followed by integration over  $X$  and  $M$  will give rise to the modified Jarzynski Equality, eq. (5).

### B. Using feedback to generate reverse process

The reverse process can also be defined by designing a suitable feedback procedure to generate the time-reversed protocol, which does not violate causality [23]. The feedback procedure is as follows: we first measure system observable at time  $t = \tau$ , and if the measurement outcome is equal to the forward outcome at time  $t = t_N$ , then we drive the system using the reverse protocol  $\lambda(m_{t_N})$  (note that the time instants for measurement along the backward process are not the same as the forward ones, but are shifted). Continuing with this process, i.e., changing the protocol according to measurement outcome  $m(t_{k-1})$  at time  $t = t_k$ , the reverse protocol is exactly reproduced without ever violating causality. This the new relation

$$\frac{P[X, M]}{\tilde{P}[\tilde{X}; M]} = e^{\beta W_d[X] + \Delta s_p[X, M]}, \quad (9)$$

where

$$\Delta s_p[X, M] \equiv \ln \frac{p(m_0|x_0)p(m_1|x_1) \cdots p(m_N|x_N)}{p(m_N|\tilde{x}_{N+1})p(m_{N-1}|\tilde{x}_N) \cdots p(m_0|\tilde{x}_1)}. \quad (10)$$

Here,  $\Delta s_p$  represents a disorder parameter, but is not the mutual information as defined in [12]. Of course, equations (8) and (9) contain different information as well as provide different bounds for the average total entropy change for the system.

### C. The most general case

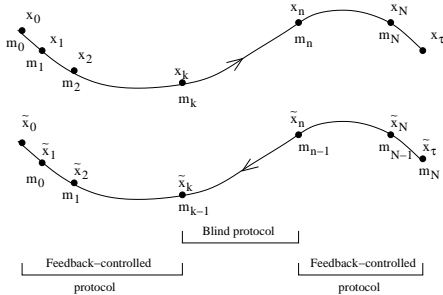


FIG. 1: The figure shows the representative plot of the phase space trajectory in a simple case for the most general form of feedback that can be considered during the reverse process (the lower curve) corresponding to the forward process (the upper curve). First, from  $t = \tau$  to  $t = t_n$ , we have the feedback-controlled reverse protocol, while from  $t = t_n$  to  $t = t_k$  we have used the blindly time-reversed protocol. In the final part, up to  $t = 0$ , we revert to the feedback-controlled protocol.

The general form of the probability density for a forward trajectory in presence of feedback is given by [16]

$$P[X, M] = p(m_0|x_0)P_{m_0}[x_0 \rightarrow x_1]p(m_1|x_1)P_{m_1}[x_1 \rightarrow x_2] \cdots p(m_N|x_N)P_{m_N}[x_N \rightarrow x_\tau]. \quad (11)$$

The general protocol to generate the reverse process would be to use both the protocols II A and II B at random during the reverse process. Let us take one simple case when, along the reverse process, up to time  $t = \tau - t_n$ , we use the feedback-controlled protocol, then from time  $t = t_n$  to  $t = t_k$ , we use the blindly applied reverse protocol, and finally from  $t = t_k$  to  $t = 0$ , we once again use the feedback-controlled reverse protocol (see figure 1). Here,  $k$  and  $n$  can be any two integers chosen at random from the set  $\{1, 2, \dots, N\}$ . In this case, the reverse process becomes,

$$\begin{aligned} \tilde{P}[\tilde{X}; M] &= p(\tilde{x}_\tau)p(m_N|\tilde{x}_\tau)P_{m_N}[\tilde{x}_\tau \rightarrow \tilde{x}_N] \\ &\quad p(m_{N-1}|\tilde{x}_N)P_{m_{N-1}}[\tilde{x}_N \rightarrow \tilde{x}_{N-1}] \\ &\cdots p(m_n|\tilde{x}_{n+1})P_{m_n}[\tilde{x}_{n+1} \rightarrow \tilde{x}_n] \\ &\quad p(m_{n-1}, \dots, m_k)P[\tilde{x}_n \rightarrow \tilde{x}_k]p(m_{k-1}|\tilde{x}_k) \\ &\quad P[\tilde{x}_k \rightarrow \tilde{x}_{k-1}] \cdots p(m_0|\tilde{x}_1)P[\tilde{x}_1 \rightarrow \tilde{x}_0]. \end{aligned} \quad (12)$$

Therefore, we arrive at

$$\begin{aligned} \frac{P[X, M]}{\tilde{P}[\tilde{X}; M]} &= \frac{P[X]}{\tilde{P}[\tilde{X}]} \frac{p(m_k|x_k) \cdots p(m_{n-1}|x_{n-1})}{p(m_k, \dots, m_{n-1})} \\ &\quad \times \frac{p(m_0|x_0) \cdots p(m_{k-1}|x_{k-1})}{p(m_0|\tilde{x}_1) \cdots p(m_{k-1}|\tilde{x}_k)} \\ &\quad \times \frac{p(m_n|x_n) \cdots p(m_N|x_N)}{p(m_n|\tilde{x}_{n+1}) \cdots p(m_N|\tilde{x}_\tau)} \\ &= \exp [\beta W_d + I^1 + \Delta s_p^1 + \Delta s_p^2], \end{aligned} \quad (13)$$

where

$$I^1 = \ln \frac{p(m_k|x_k) \cdots p(m_{n-1}|x_{n-1})}{p(m_k, \dots, m_{n-1})}; \quad (14)$$

$$\Delta s_p^1 = \ln \frac{p(m_0|x_0) \cdots p(m_{k-1}|x_{k-1})}{p(m_0|\tilde{x}_1) \cdots p(m_{k-1}|\tilde{x}_k)}; \quad (15)$$

$$\Delta s_p^2 = \ln \frac{p(m_n|x_n) \cdots p(m_N|x_N)}{p(m_n|\tilde{x}_{n+1}) \cdots p(m_N|\tilde{x}_\tau)}. \quad (16)$$

We are thus led to a different extended detailed fluctuation theorem where disorder parameters  $\Delta s_p^1$ ,  $I^1$  and  $\Delta s_p^2$  are different as they contain different information about the feedback process. Moreover, these are not unique and may depend on different reverse protocols not considered here. Thus, it is clear that the  $\phi[X, M]$  in eq. (8) does not have a unique interpretation and clearly depends on the manner of feedback along the backward process. From eq. (13), one can obtain the related integral fluctuation theorem, namely

$$\langle \exp[-\beta W_d - I^1 - \Delta s_p^1 - \Delta s_p^2] \rangle = 1 \quad (17)$$

All the three results derived above, eqs. (8), (9) and (13), can be written in compact form as

$$\frac{P[X, M]}{\tilde{P}[\tilde{X}; M]} = e^{\beta W_d + \phi[X, M]}, \quad (18)$$

where  $\phi[X, M]$  is a functional of the phase space trajectory as well as of the measurement trajectory (i.e., sequence of measured outcomes).

We next turn our attention to the efficacy parameter.

### III. EFFICACY PARAMETER IN PRESENCE OF GENERAL FEEDBACK

In absence of feedback, the Jarzynski equality is given by [3–6]

$$\langle e^{-\beta W_d[X]} \rangle = 1. \quad (19)$$

In presence of feedback, the right hand side of the above relation will in general not be unity. The efficacy parameter for the feedback in this case (when system is initially at thermal equilibrium with the medium) is defined as

$$\gamma = \langle e^{-\beta W_d[X]} \rangle. \quad (20)$$

It can be readily seen that lesser the amount of dissipated work, more is the magnitude of the efficacy parameter, as is desirable for a quantity that decides the efficiency of feedback. Now we use the most general case for obtaining the reverse trajectories, namely the case II C.

Therefore, we get

$$\begin{aligned} \gamma &= \int \mathcal{D}[X, M] P[X, M] e^{-W_d[X, M]} \\ &= \int \mathcal{D}[X, M] \tilde{P}[\tilde{X}; M] e^{\Delta s_p^1 + I^1 + \Delta s_p^2} \\ &= \int \mathcal{D}[X, M] \tilde{P}[\tilde{X}] p(m_0|x_0) p(m_1|x_1) \\ &\quad \cdots p(m_N|x_N) \\ &= \int \mathcal{D}[\tilde{X}, \tilde{M}] \tilde{P}[\tilde{X}] p(\tilde{m}_0|\tilde{x}_0) \cdots p(\tilde{m}_N|\tilde{x}_N) \\ &= \int \mathcal{D}[\tilde{X}, \tilde{M}] \tilde{P}[\tilde{X}, \tilde{M}] = \int \mathcal{D}[\tilde{M}] \tilde{P}[\tilde{M}], \end{aligned} \quad (21)$$

where we have used the detailed fluctuation theorem (18), and the definitions (15), (14) and (16). The time-reversibility of measurements has been assumed:  $p(m_i|x_i) = \tilde{p}(\tilde{m}_i|\tilde{x}_i)$  [12]. Once again we may recall that the two quantities  $\tilde{P}[\tilde{X}; M]$  and  $\tilde{P}[\tilde{X}, \tilde{M}]$ , both of which appear in the above derivation, are different. In more general cases, the steps 2 and 3 can be written as

$$\begin{aligned} \tilde{P}[\tilde{X}; M] e^{\phi[X, M]} &= \tilde{P}[\tilde{X}; M] \exp \left( \sum_i \Delta s_p^i + \sum_j I^j \right) \\ &= \tilde{P}[\tilde{X}] p(m_1|x_1) \cdots p(m_N|x_N) \\ &= \tilde{P}[\tilde{X}] \tilde{p}(\tilde{m}_1|\tilde{x}_1) \cdots \tilde{p}(\tilde{m}_N|\tilde{x}_N) \\ &= \tilde{P}[\tilde{X}, \tilde{M}]. \end{aligned} \quad (22)$$

Here, the summations  $\sum_i$  and  $\sum_j$  run over all the time intervals in which the reverse protocols have been executed using feedback and by blind time-reversal, respectively. We find that although the form of  $\tilde{P}[\tilde{X}; M]$  contains detailed information about the actual feedback procedure used along the backward process, when it is multiplied by the factor  $e^{\phi[X, M]}$ , we obtain  $\tilde{P}[\tilde{X}, \tilde{M}]$  whose form does not contain any such information. This is the

reason behind the fact that the efficacy parameter retains the same physical meaning in each case, namely, it is the *total probability to observe the time-reversed outcomes for the measurements performed along the reverse process* [11–13, 16].

This derivation of  $\gamma$  for the most general case is our central result. It simply shows that it retains the same meaning as for the specific cases considered in the earlier literature [12–19].

### IV. THE THREE DETAILED FLUCTUATION THEOREMS

We will be generalizing our treatment to the other detailed fluctuation theorems, which involve the non-adiabatic entropy production (change in entropy due to shift between two different steady states) and adiabatic entropy production (change in entropy caused due to maintenance of a given steady state). The total entropy change by definition is the sum of entropy changes in the system ( $\Delta s$ ) and in the medium ( $\Delta s_m$ ):  $\Delta s_{tot} = \Delta s + \Delta s_m$ . Recently it has been observed that while generalizing the second law for systems making transitions between steady states, the total entropy production can also be split into two distinct parts such that each part, interestingly, follows a detailed fluctuation theorem [24–27]:

$$\Delta s_{tot} = \Delta s_{na} + \Delta s_a. \quad (23)$$

The averages of all these three quantities are always non-negative, thereby providing a new twist to the second law.  $\Delta s_a$  is related to the housekeeping heat  $Q_{hk}$ , while  $\Delta s_{na}$  is the sum of the entropy change of the system and the entropy produced due to excess heat  $Q_{ex}$  [24–27].

In the case of adiabatic and non-adiabatic entropy productions, the concept of dual dynamics is very helpful. Under the dual dynamics, if the system is allowed to reach the corresponding steady state, then the steady-state distribution  $\rho_{SS}$  retains the same form, but the probability current reverses its sign [27, 28]. Hatano and Sasa had shown that the physical meaning of the nonadiabatic entropy becomes clear in the dual dynamics formalism [29]. These detailed fluctuation theorems are taken up in the following discussion.

#### A. Total entropy

When the initial state of the system is not at thermal equilibrium with the bath, then in absence of feedback, the following ratio is obtained between the forward and the reverse trajectories [7, 8, 10]:

$$\frac{P[X]}{\tilde{P}[\tilde{X}]} = e^{\Delta s_{tot}[X]}, \quad (24)$$

from which the following integral fluctuation theorem can be obtained:

$$\langle e^{-\Delta s_{tot}[X]} \rangle = 1. \quad (25)$$

In presence of feedback, the right hand side will in general be different from unity. For this general case, instead of eq. (18), we would get the following ratio between the forward and reverse paths:

$$\frac{P[X, M]}{\tilde{P}[\tilde{X}; M]} = e^{\Delta s_{tot}[X] + \phi[X, M]}. \quad (26)$$

We now consider the case with general reverse protocol. We define the efficacy parameter as

$$\gamma_{tot} = \langle e^{-\Delta s_{tot}[X]} \rangle \quad (27)$$

Proceeding in exactly the same way as before, we find

$$\begin{aligned} \gamma_{tot} &= \int \mathcal{D}[X, M] P[X, M] e^{-\Delta s_{tot}[X, M]} \\ &= \int \mathcal{D}[X, M] \tilde{P}[\tilde{X}; M] e^{\phi[X, M]} \\ &= \int \mathcal{D}[\tilde{M}] \tilde{P}[\tilde{M}]. \end{aligned} \quad (28)$$

Thus,  $\gamma_{tot}$  retains the same physical meaning as  $\gamma$  for the Jarzynski equality, although here we do not have the constraint of sampling the initial state of the system from the Boltzmann distribution.

### B. Nonadiabatic entropy

For transitions between nonequilibrium steady states, we have the following detailed fluctuation theorem in absence of feedback [25, 27, 29]:

$$\frac{P[X]}{\tilde{P}^\dagger[\tilde{X}]} = e^{\Delta s_{na}[X]}. \quad (29)$$

The superscript  $\dagger$  implies dual dynamics, which implies that keeping the functional form of the protocol same, we are switching to a dynamics where the steady state current simply changes direction as compared to the original dynamics. The tilde symbol over  $P$  implies that the protocol for the forward process has been time-reversed after the system has been allowed to follow the dual dynamics. In other words,  $\tilde{P}^\dagger[\tilde{X}]$  is the probability density for a trajectory along the process generated, in presence of dual dynamics, by the time-reversed protocol. Similar to the above cases, in presence of feedback, we have [16, 19]

$$\frac{P[X, M]}{\tilde{P}^\dagger[\tilde{X}; M]} = e^{\Delta s_{na}[X] + \phi[X, M]}, \quad (30)$$

where the form of  $\phi[X, M]$  depends on the way in which feedback is applied along the reverse trajectory, as given

in section 2. The efficacy parameter in this case is given by

$$\begin{aligned} \gamma_{na} &= \langle e^{-\Delta s_{na}[X]} \rangle = \int \mathcal{D}[X, M] P[X, M] e^{-\Delta s_{na}[X, M]} \\ &= \int \mathcal{D}[X, M] \tilde{P}^\dagger[\tilde{X}; M] e^{\phi[X, M]} \\ &= \int \mathcal{D}[X, M] \tilde{P}^\dagger[\tilde{X}] p(m_1|x_1) \cdots p(m_N|x_N) \\ &= \int \mathcal{D}[X, M] \tilde{P}^\dagger[\tilde{X}, \tilde{M}] = \int \mathcal{D}[\tilde{M}] \tilde{P}^\dagger[\tilde{M}]. \end{aligned} \quad (31)$$

In the third step, we have used the algebra that has already been shown in the case of the most general protocol (section III) for extended Jarzynski equality, which results in

$$\begin{aligned} \tilde{P}^\dagger[\tilde{X}; M] e^{\phi[X, M]} &= \tilde{P}^\dagger[\tilde{X}] \exp \left( \sum_i \Delta s_p^i + \sum_j I^j \right) \\ &= \tilde{P}^\dagger[\tilde{X}] p(m_1|x_1) \cdots p(m_N|x_N) \\ &= \tilde{P}^\dagger[\tilde{X}] \tilde{p}^\dagger(\tilde{m}_1|\tilde{x}_1) \cdots \tilde{p}^\dagger(\tilde{m}_N|\tilde{x}_N) \\ &= \tilde{P}^\dagger[\tilde{X}, \tilde{M}]. \end{aligned} \quad (32)$$

We have assumed that the measurement errors do not change on changing the dynamics, which is quite reasonable assumption. Thus,  $\gamma_{na}$  is the net probability for obtaining the time-reversed outcomes along the time-reversed dual dynamics (i.e., the process in which the protocol corresponding to the dual dynamics is time-reversed).

### C. Adiabatic entropy

The DFT for adiabatic entropy production is given by [25, 27]

$$\frac{P[X]}{\tilde{P}^\dagger[\tilde{X}]} = e^{\Delta s_a[X]}. \quad (33)$$

$\tilde{P}^\dagger[\tilde{X}]$  is the probability density for the path followed by the system in phase space, when the system is evolving under the dual dynamics. In presence of feedback, we then have,

$$\frac{P[X, M]}{\tilde{P}^\dagger[\tilde{X}; M]} = e^{\Delta s_a[X] + \phi[X, M]}. \quad (34)$$

Since there is no time-reversal involved, the denominator can only consist of the following options:

(1) The same feedback procedure is used to generate the forward process in the dual dynamics as well, in which case we have  $\phi[X, M] = 0$  (since the error probabilities in the numerator cancel with those in the denominator). Once again, measurement errors are assumed to be independent of the dynamics followed by the system.

(2) One of the forward protocols in the original dynamics is recorded, and this protocol is blindly executed in presence of the dual dynamics, in which case we simply have  $\phi[X, M] = I[X, M]$ .

(3) We use the above two procedures at random while generating the forward trajectories in presence of dual dynamics, which is the most general case. In this case, however,  $\phi[X, M] = \sum_j I^j$ , i.e., the summation over  $\Delta s_p^i$  will be absent, because the latter quantity never appears in this case.

The efficacy parameter is

$$\gamma_a \equiv \langle e^{-\Delta s_a[X]} \rangle, \quad (35)$$

which leads to

$$\begin{aligned} \gamma_a &= \int \mathcal{D}[X, M] P[X, M] e^{-\Delta s_a[X, M]} \\ &= \int \mathcal{D}[X, M] P^\dagger[X; M] e^{\sum_j I^j} = \int \mathcal{D}[M] P^\dagger[M]. \end{aligned} \quad (36)$$

Therefore,  $\gamma_a$  is the total probability for observing the same outcomes as the initial process with the same protocol, if the system follows the dual dynamics.

We thus find that the physical meaning of efficacy parameter can be very generally stated as follows: *it is the total probability to observe the measured outcomes conjugate to those along the forward protocol, for the intermediate measurements along the process with the corresponding conjugate dynamics.* Since the efficacy parameters are experimentally measurable, they would provide more meaningful bounds for the  $\langle W_d \rangle$ ,  $\langle \Delta s_a \rangle$ ,  $\langle \Delta s_{na} \rangle$ ,  $\langle \Delta s_{tot} \rangle$ , and these bounds are in fact independent of whether or not feedback is performed along the conjugate process in the actual protocol, but only requires measurements to be performed along the conjugate process. For the other bounds stated in section 2, the values would depend sensitively on whether and how the feedback is performed along the conjugate dynamics, namely, the extended integral fluctuation theorems can be stated as

$$\left\langle \exp \left( -\Delta s_k - \sum_i \Delta s_p^i - \sum_j I^j \right) \right\rangle = 1 \quad (37)$$

for  $k = 1, 2, 3$  representing  $\Delta s_{tot}$ ,  $\Delta s_{na}$  and  $\Delta s_a$ , respectively, and this would lead to the bounds

$$\langle \Delta s_k \rangle \geq - \left\langle \sum_i \Delta s_p^i \right\rangle - \left\langle \sum_j I^j \right\rangle. \quad (38)$$

As a consequence, arbitrary number of bounds can be computed for this latter case, which is not only confusing, but undermines the motivation behind such calculations. Thus, the efficacy parameter is a far more suitable experimentally measurable quantity that can characterize not only the performance of the system, but can also act as a useful lower bound for the averaged observables.

## V. CONCLUSIONS

In this paper, we have shown that out of the two known forms of the modified fluctuation theorems in presence of feedback, one of the forms is heavily dependent on the way feedback is applied along the conjugate process, and thereby leads to arbitrary number of lower bounds for work done on the system or for the relevant entropy changes (total entropy, nonadiabatic and adiabatic entropies) taking place. On the other hand, the second form, namely the fluctuation theorem expressed in terms of the efficacy parameter, provides a bound for work and entropy changes that carries a clear and consistent physical meaning, irrespective of the manner of application of feedback along the conjugate process. This consistency is robust even when the conjugate process is not the time-reversed process. This study would hopefully help in simpler experimental verification of the extended fluctuation theorems.

## VI. ACKNOWLEDGEMENT

One of us (AMJ) thanks DST, India for financial support. We thank Dr. Anupam Kundu for sending us the preprint of his manuscript prior to its publication.

---

[1] D. J. Evans, E. G. D. Cohen and G. P. Morriss, Phys. Rev. Lett. **71**, 2401 (1993).  
[2] D. J. Evans and D. J. Searles, Phys. Rev. E **50**, 1645 (1994).  
[3] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).  
[4] C. Jarzynski, Phys. Rev. E **56**, 5018 (1997).  
[5] G. E. Crooks, J. Stat. Phys. **90**, 1481 (1998).  
[6] G. E. Crooks, Phys. Rev. E **60**, 2721 (1999).  
[7] U. Seifert, Phys. Rev. Lett. **95**, 040602 (2005).  
[8] U. Seifert, Eur. Phys. J. B **64**, 423 (2008).  
[9] M. Campisi, P. Hänggi and P. Talkner, Rev. Mod. Phys. **83**, 771 (2011).

[10] U. Seifert, arXiv:1205.4176.  
[11] T. Sagawa and M. Ueda, Phys. Rev. Lett. **100**, 080403 (2008).  
[12] T. Sagawa and M. Ueda, Phys. Rev. Lett. **104**, 090602 (2010).  
[13] T. Sagawa and M. Ueda, Phys. Rev. E **85**, 021104 (2012).  
[14] J. M. Horowitz and S. Vaikuntanthan, Phys. Rev. E **82**, 061120 (2010).  
[15] M. Ponmurugan, Phys. Rev. E **83**, 031129 (2010).  
[16] S. Lahiri, S. Rana and A. M. Jayannavar, J. Phys. A: **44**, 115001 (2011).

- Math. Theor. **45**, 065002 (2012).
- [17] S. Rana, S. Lahiri and A. M. Jayannavar, arXiv:1202.1097.
  - [18] S. Rana, S. Lahiri and A. M. Jayannavar, arXiv:1111.6229; accepted for publication in Pramana J. Phys.
  - [19] D. Abrieu and U. Seifert, Phys. Rev. Lett. **108**, 030601 (2012).
  - [20] F. J. Cao and M. Feito, Phys. Rev. E **79**, 041118 (2009).
  - [21] T. M. Cover and J. A. Thomas, *Elements of information theory*, 2nd edn (Hoboken, NJ: Wiley-Interscience).
  - [22] S. Toyabe et al, Nature Physics **6**, 988 (2010).
  - [23] A. Kundu, private communication.
  - [24] M. Esposito and C. van den Broeck, Phys. Rev. Lett. **104**, 090601 (2010).
  - [25] M. Esposito and C. van den Broeck, Phys. Rev. E **82**, 011143 (2010).
  - [26] M. Esposito and C. van den Broeck, Phys. Rev. E **82**, 011144 (2010).
  - [27] R. Garcia-Garcia, V. Lecomte, A. B. Kolton and D. Dominguez, J. Stat. Mech. P02009 (2012).
  - [28] V. Y. Chernayk, M. Chertkov and C. Jarzynski, J. Stat. Mech. P08001 (2006).
  - [29] T. Hatano and S. Sasa, Phys. Rev. Lett. **86**, 3463 (2001).